OPTIMIZING MANUFACTURING SUPPLY CHAIN

Purdue Student Labs

PURDUE UNIVERSITY
Mitchell E. Daniels, Jr.
School of Business

AGENDA

- Objective
- Data & Methodology
- Modeling
- Results
- Conclusion & Future Scope

PURDUE UNIVERSITY | Mitchell E. Daniels, Jr. School of Business

OBJECTIVE

STAKEHOLDER INTRODUCTION

Introduction

- Leading name in the welding industry for the last 129 years
- Offers a diverse range of **Welding Products** including Welding Machines, Consumables, Cutting Equipment, Automation Systems, and Accessories.

Manufacturing & Supply Chain

- Extensive product variety: **1,304 finished goods.**
- Global raw material procurement from over **800 suppliers**.
- Emphasizes strong, long-term supplier relationships.

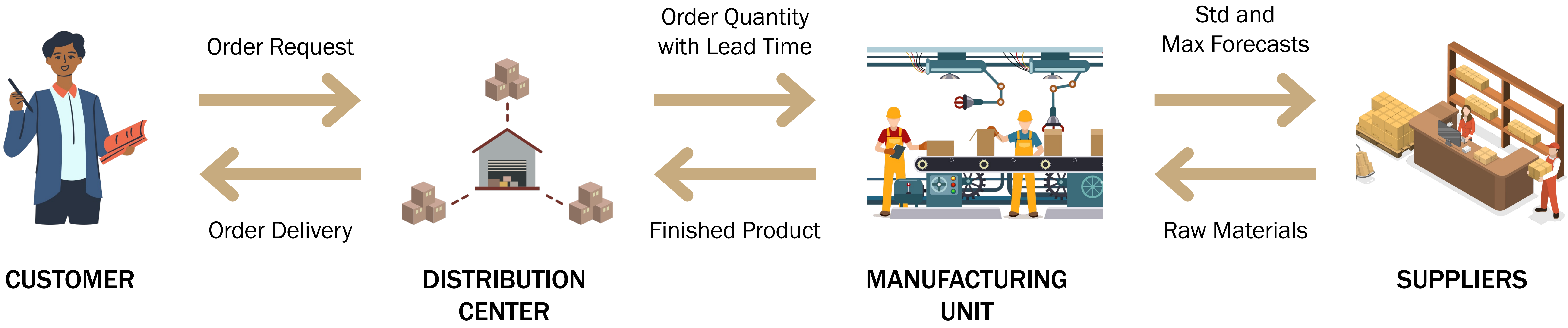
Production Approach

- Utilizes a **Pull Mechanism** to align production closely with demand.
- Adheres to Lean Manufacturing principles to optimize efficiency and minimize waste.

US-based welding equipment manufacturer

BUSINESS PROBLEM

Impact of Inaccurate Demand Forecasting on Supply Chain



Inaccurate demand forecasting amid market uncertainties severely impacts a complex supply chain:

1. Impact on Firm: Inaccurate forecasts lead to stockouts, impacting production efficiency, customer satisfaction, and financial health of manufacturing firms.
2. Impact on Suppliers: Inaccurate forecasts lead to inventory surplus, tying up financial resources, causing financial strain and potentially damaging long-term business relationships within the industry.

PROBLEM STATEMENT

Key Challenges

- **Traditional Forecasting Techniques:**
  Reliance on historical averages for predicting demand of materials with erratic demand patterns and minimal correlation.
- **Overestimation of Inventory:**
  While actual quantity ordered aligns closer to standard, max forecasts are double the standard forecasts, causing supplier discontent due to surplus.

Project Scope

- Develop advanced machine learning and time-series predictive models to address complexities of strategic products
- Improved forecasting algorithm to reduce the variation between Standard and Max forecasts of the final product
- Service parts and new products are excluded

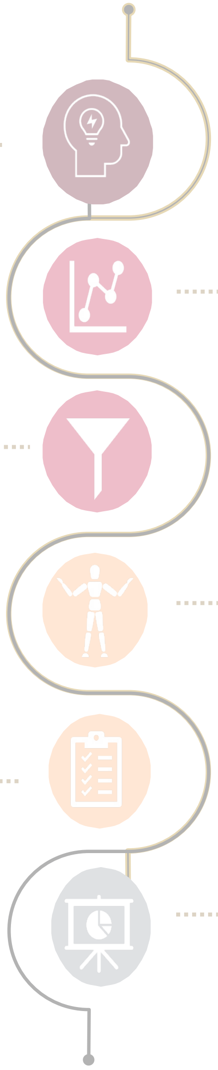
Key Deliverables

- Forecasting framework to provide monthly forecasts including Jupyter notebooks and output files
- Tableau report of the forecasts at at material – month level
- Knowledge transfer document with requirements and steps to replicate the analysis and produce results

PROJECT OVERVIEW

Developed and implemented an advanced machine learning solution to **accurately forecast the volatile demand for materials** in the welding industry, overcoming the limitations of traditional forecasting methods.

Methodology



Data Understanding
- Entity Relationship Diagrams
- Data Dictionaries

Data Preprocessing
- Outlier & Missing Value Treatment
- Master Dataset Creation
- Data Transformation and Product Segmentation

Exploratory Data Analysis (EDA)
- Product Segmentation
- Time Series Decomposition
- Graphical Visualization

Data Modelling & Validation

Create time series forecast, Evaluate Model Outcomes, to get Best Fit using Dynamic Selection Box:
- Holt Winter
- ARIMA
- Exponential Smoothing
- Prophet

Reporting & Insights
- Tableau Dashboard
- Recommendations Report

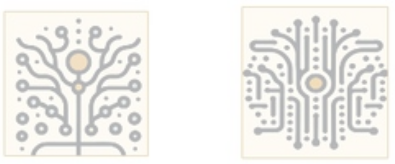
Forecast Extension
- Explode the finished product forecasts to raw material level

Modeling

Advanced Time series models
1. Low Volume – Low Variance
2. Low Volume – High Variance

SARIMA Holt Winter Prophet

Advanced ML models
1. High Volume – Low Variance
2. High Volume – High Variance



Wave Net CNN

Dynamic Selection box

Key Metrics: Use lowest WMAPE to select the best model for each material

$$WMAPE = \frac{1}{\sum_{i=1}^{n} A_i} \sum_{i=1}^{n} | A_i - F_i$$

- A_i as the actual quantity for month i
- P_i as the predicted quantity for month i
- n is the total number of months considered

Model Outputs

Standard Forecast and Standard deviation

Calculating Model Improvement Factor (MIF)

$$\frac{\text{Abs[Avg. WMAPE}_{\text{Current}} - \text{Avg. WMAPE}_{\text{New}}\text{]}}{\text{Avg. WMAPE}_{\text{Current}}}$$

Calculating Maximum Forecast

Max{Std. Forecast + 1.65*Std. Deviation * (1-MIF), 2* Std. Forecast }

Steps to calculate Max. Forecast
1. Calculate MIF, which capture improvement factor of the new model over current model
2. Apply MIF to the 90% Confidence Interval (CI) estimate of Std. forecast

Final Output

Standard Forecast and Max. Forecast (90% CI)

Results

$$WMAPE = \frac{1}{\sum_{i=1}^{n} A_i} \sum_{i=1}^{n} |A_i - F_i|$$

Weighted MAPE has been used to compare the accuracy of forecasts before and after the framework was implemented

New WMAPE: 43%

Old WMAPE: 60%

This indicates an improvement of 17 **percentage points** at an overall level

DATA & METHODOLOGY

SOLUTION APPROACH

Five-Step Data Analytics Lifecycle

Step One

Data Understanding & Exploratory Data Analysis

- Create Data Dictionaries
- Entity Relationship Diagrams
- Graphical Visualization

Step Two

Data Preprocessing

- Outlier & Missing Value Treatment
- Data Transformation & Master Dataset Creation
- Volume by Variance Product Segmentation

Step Three

Data Modelling & Validation

- Add Cyclic Features
- Dynamic Selection Box
- Time Series & ML Models

Step Four

Reducing Max Forecast

- Calculate Model Improvement Factor
- Calculate Max Forecast

Step Five

Reporting & Recommendations

- Tableau Dashboard
- Forecasting Report
- KT Document

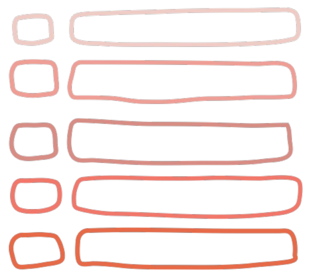
DATASET INFORMATION

Key Columns

Vital columns for Forecast:
- Order Request Date
- Material ID
- Quantity Ordered

Additional tables for expansion to raw material level:
- Product Hierarchy
- Component Parts
- Supplier’s Delivery Performance

Key data variables:
Purchase ID, Product ID, Quantity Ordered, Order Request Date, Order Delivery Date and In-House Production Time



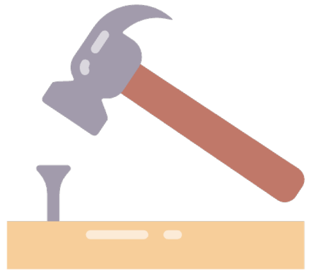
Total No. of Rows & Columns
358K Rows & 19 Columns

Materials Rows & Relevant Columns
148K Rows

Time Period (Requisition date)
06/04/2014 – 12/22/2023

Time Period (Delivery date)
01/15/2014 – 04/18/2024

Total Quantity Requested
350+



Total No. of Distinct Materials
1304

PRODUCT SEGMENTATION

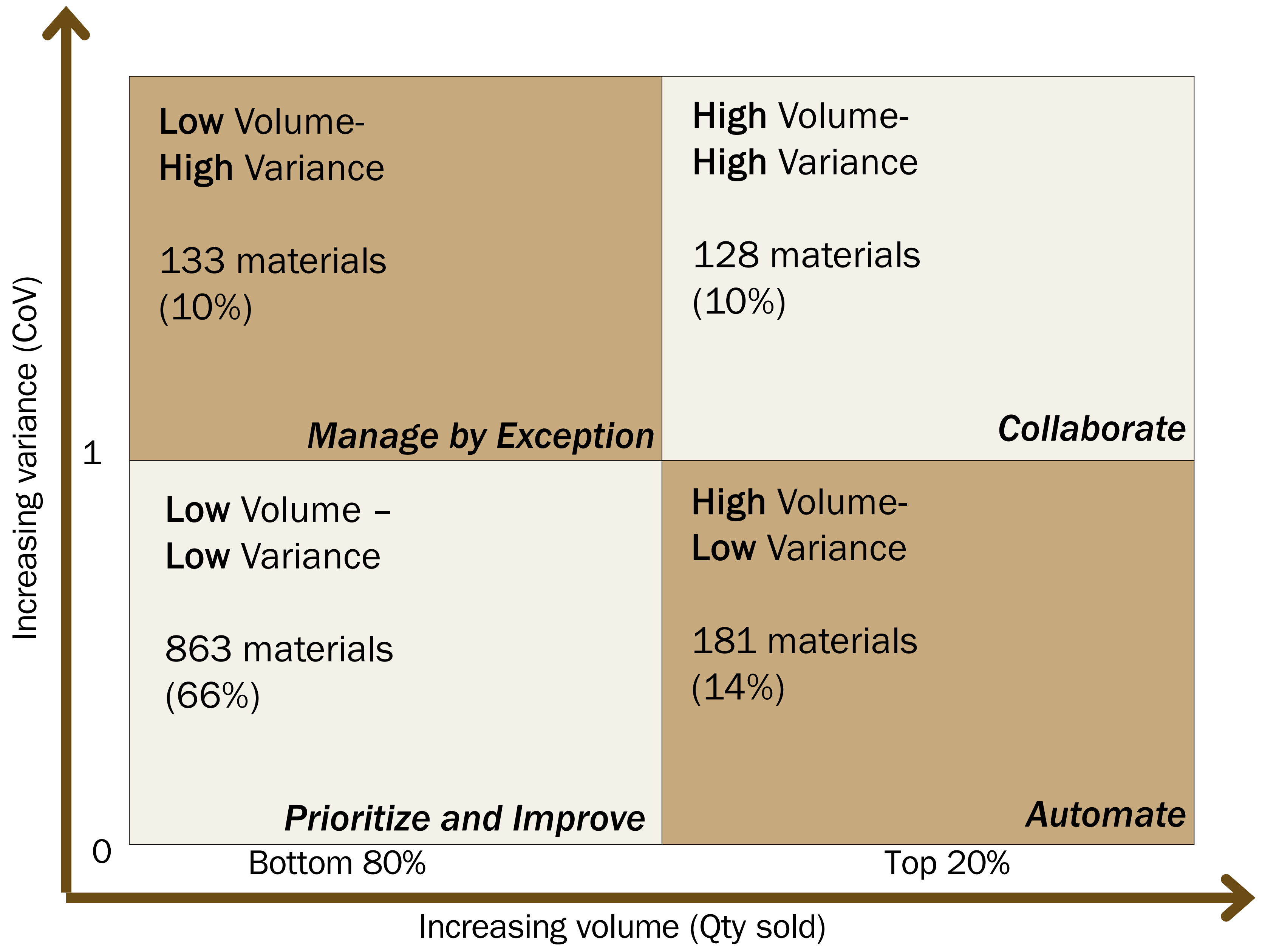
1304 finished products have been segmented into 4 categories to account for variability in demand.
This segmentation strategy is crucial for developing customized time-series and ML models considering mixed nature of products' demand

Segment description for Variance

- High Variance = Coefficient of Variation > 1
- Low Variance = Coefficient of Variation <= 1
- Coefficient of Variation is degree of variation in data defined as ratio of std deviation to average

Segment description for volume

- High Volume = Top 20% materials by qty sold
- Low Volume = Bottom 80% materials by qty sold



MODELING

DATA MODELING

Forecast Improvement & Max Reduction

Machine Learning Approach for Low Volume

Deploy classic ML models such as ARIMA and Prophet to make forecasts

Deep Learning Approach for High Volume

Classification model to predict if there will be demand for a day

Forecasting model to predict demand for days with demand > 0

MIF is a measure of the improvement in new standard forecasts when compared to the existing standard forecasts

Dynamic Selection box: Use lowest WMAPE to select the best model for each material

Derive monthly forecast and standard deviation values for each material

Calculate **Model Improvement Factor**

Calculate **Max. Forecast**

$$WMAPE = \frac{1}{\sum_{i=1}^{n} A_i} \sum_{i=1}^{n} |A_i - F_i|$$

$$MIF = \frac{Abs\ [(CM - NM)]}{CM}$$

*Max{ SF + 1.65 * SD * (1-MIF) , 2* SF }*

The model computes the standard forecast at a 90% confidence interval and uses the MIF to calculate the adjusted final maximum forecast

PURDUE UNIVERSITY | Mitchell E. Daniels, Jr. School of Business

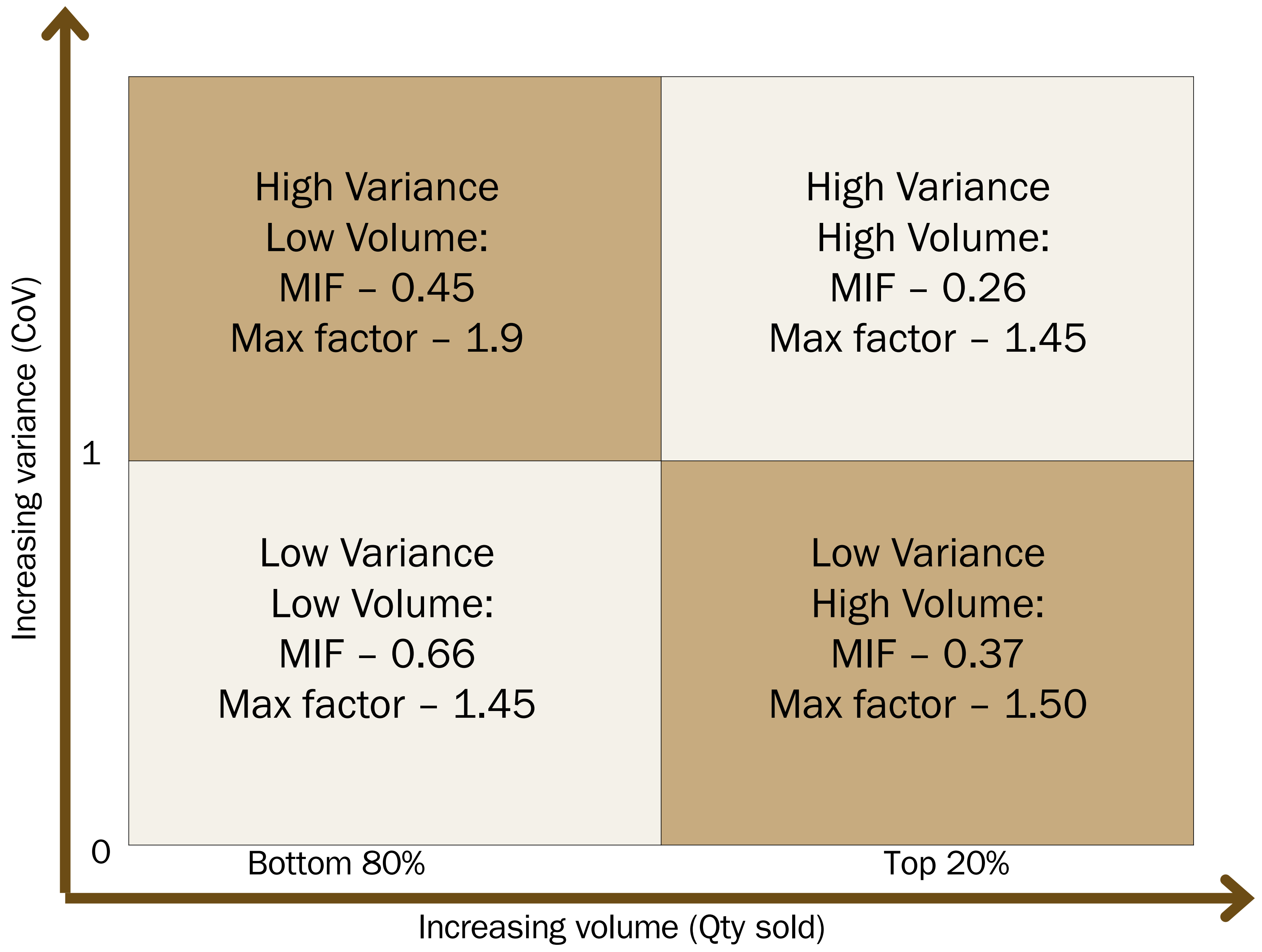
*CM - Current WMAPE
*NM - New WMAPE
*SF - Standard Forecast
*SD - Standard Deviation
*MIF - Improvement Factor

VALIDATION METRICS

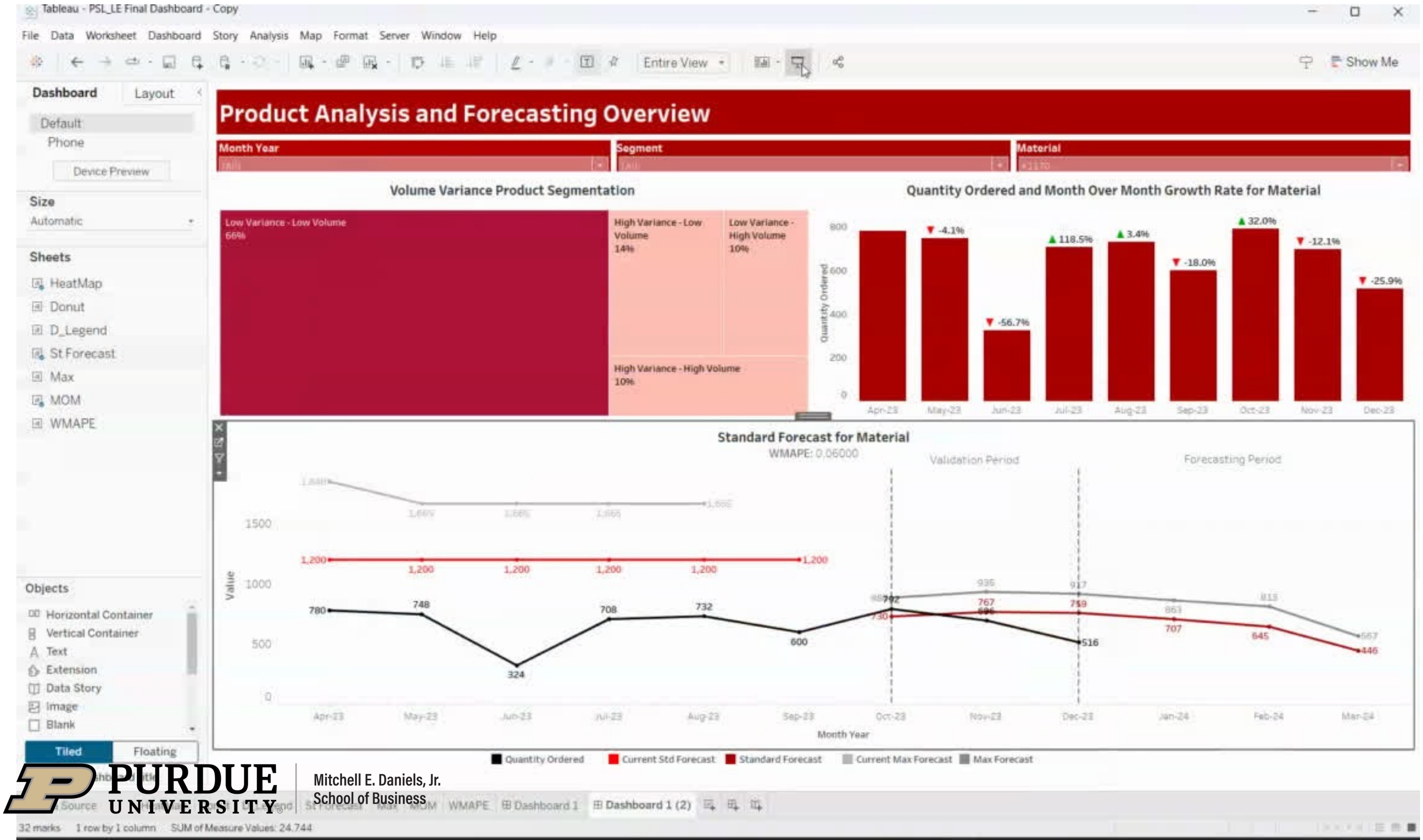
Measuring forecast improvement metrics.

- **MIF:** used to gauge the **improvement in standard forecasts**
- **Max Factor:** values for the new and existing values to gauge the **improvement in max forecasts**

- Max Factor equals
  Max Forecast / Standard Forecast
- The objective is to reduce it when compared to **Old Max Factor of 2**



RESULTS

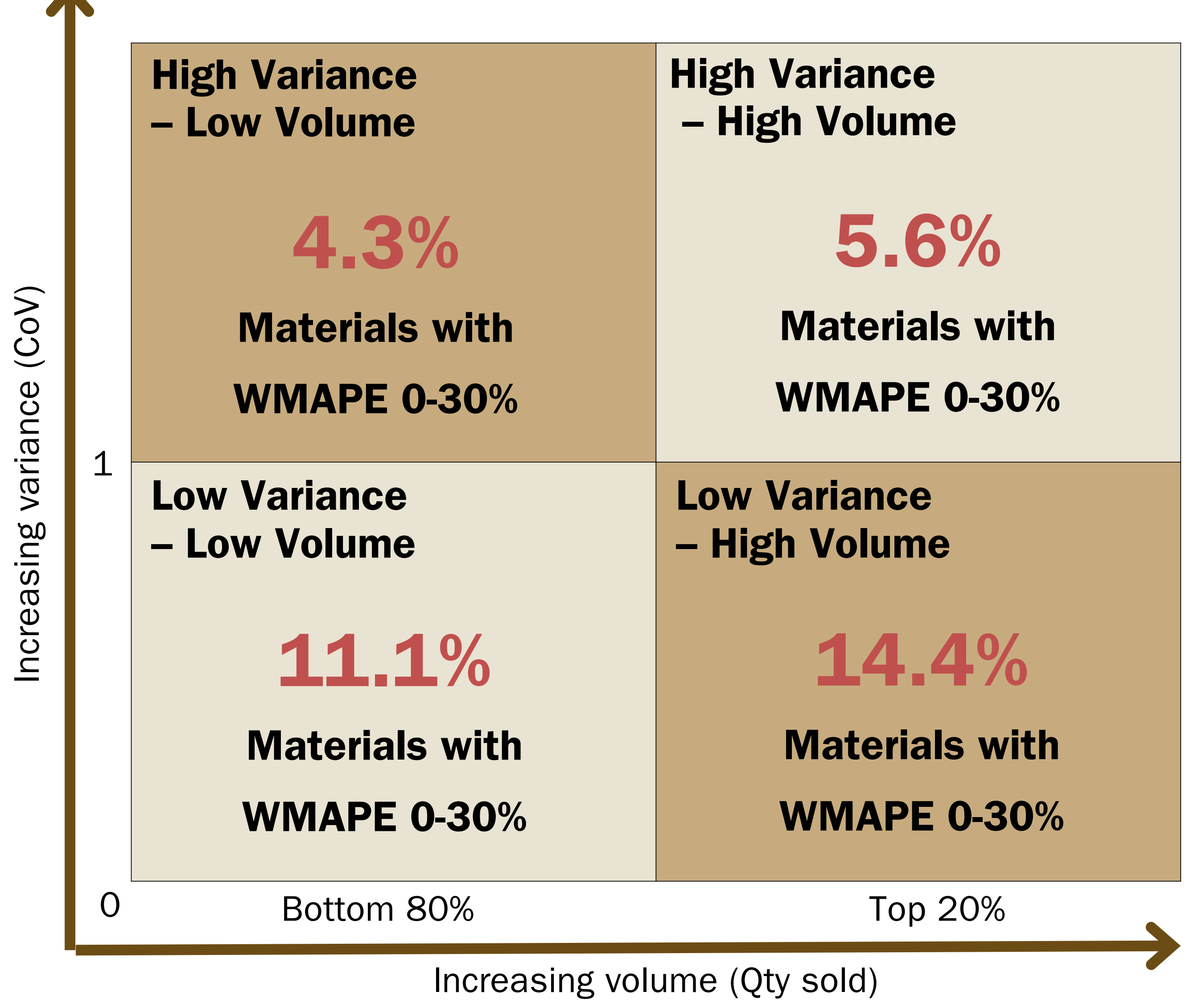


Mitchell E. Daniels, Jr.
School of Business

RESULT COMPARISON

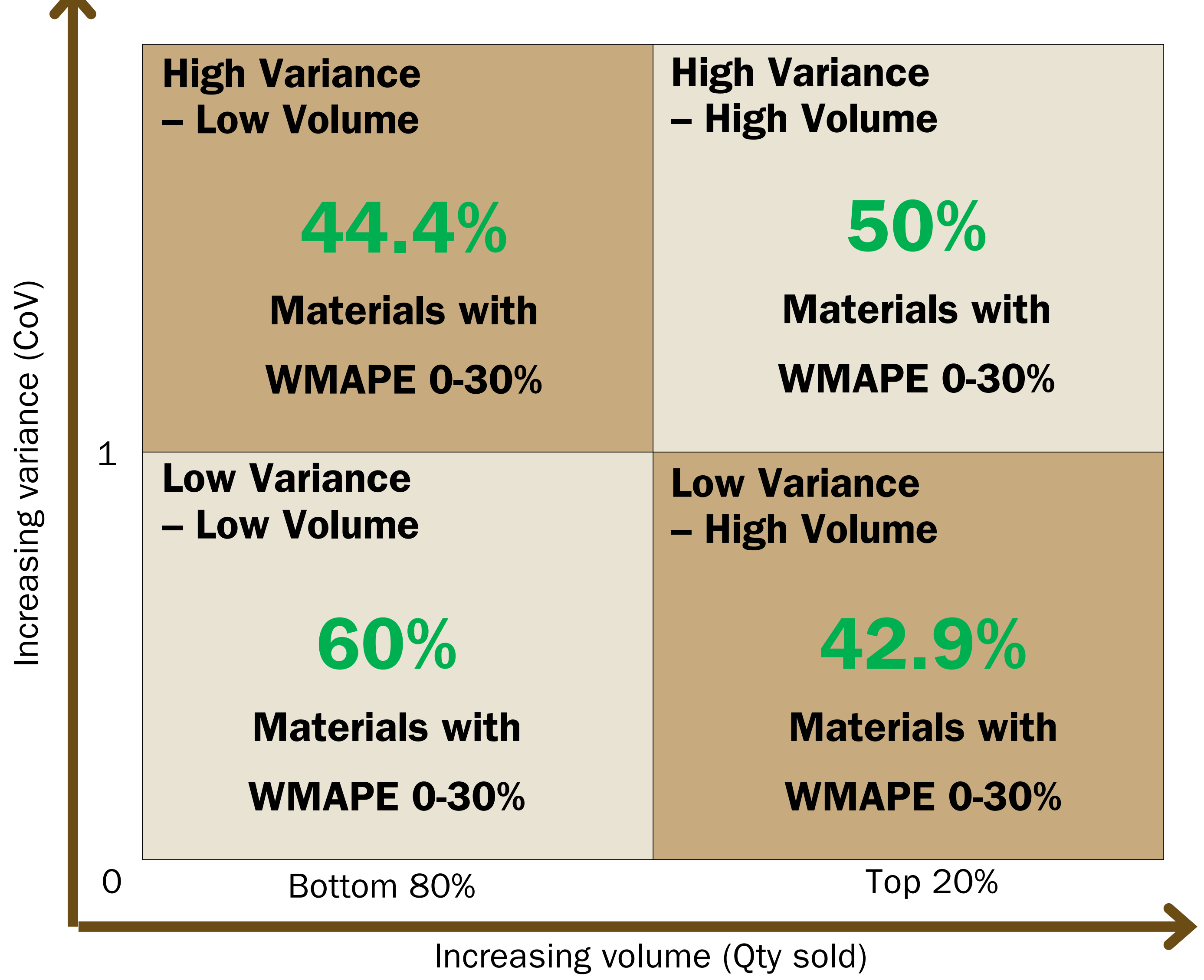
The comparison shows significant improvement in all categories, with the most notable changes being a drastic increase from 4.3% to 44.4% in the 'High Variance – Low Volume' and from 14.4% to 42.9% in the 'Low Variance – High Volume' segments for materials within the WMAPE 0-30% range.

Existing results



VS

New results



KEY HIGHLIGHTS, CONCLUSION & FUTURE SCOPE

KEY HIGHLIGHTS AND CONCLUSION

01

Inclusion of Classification in the model set up helped increase improvement in MAPE by 10%

Classification Approach

02

Weighted MAPE improvement in all segments with at least 17% improvement across all segments

WMAPE Improvement

03

Best fit model approach enables dynamic forecasting, thus allowing multiple models to fit for each material

Best Fit Model

04

MIF helped reduce the Max factor significantly, with maximum materials having a max factor between 1-1.5

MIF Factor

FUTURE SCOPE

Additional Data Sources to be Used

Stock Out Data	Historical order data can be enriched by adding information about lost sales and stock outs
Product Supersession	Addition of product life stage to the data as qualitative information will further enable accurate forecasting methodology
Longer Timeframe	The analysis was limited to data starting 2021. Extend the training dataset to improved identification of troughs and crests
Supplier Data	Information about supplier lead times will help paint the entire picture and tie back the forecasts for improved planning of purchases

THANK YOU!

MEET THE TEAM!

Akanksha Singh

Data science consultant with four years of analytical experience in retail, sports and home improvement industry

Chaitanya Krishna Burri

IT Project Manager with 9 years of experience in Telecom, Banking and Financial Services sectors

Priya Sharma

Business Analyst with three years of experience in banking sector

Sathwik Kanukuntla

Data-driven Consultant with three years of experience in manufacturing and IT sectors

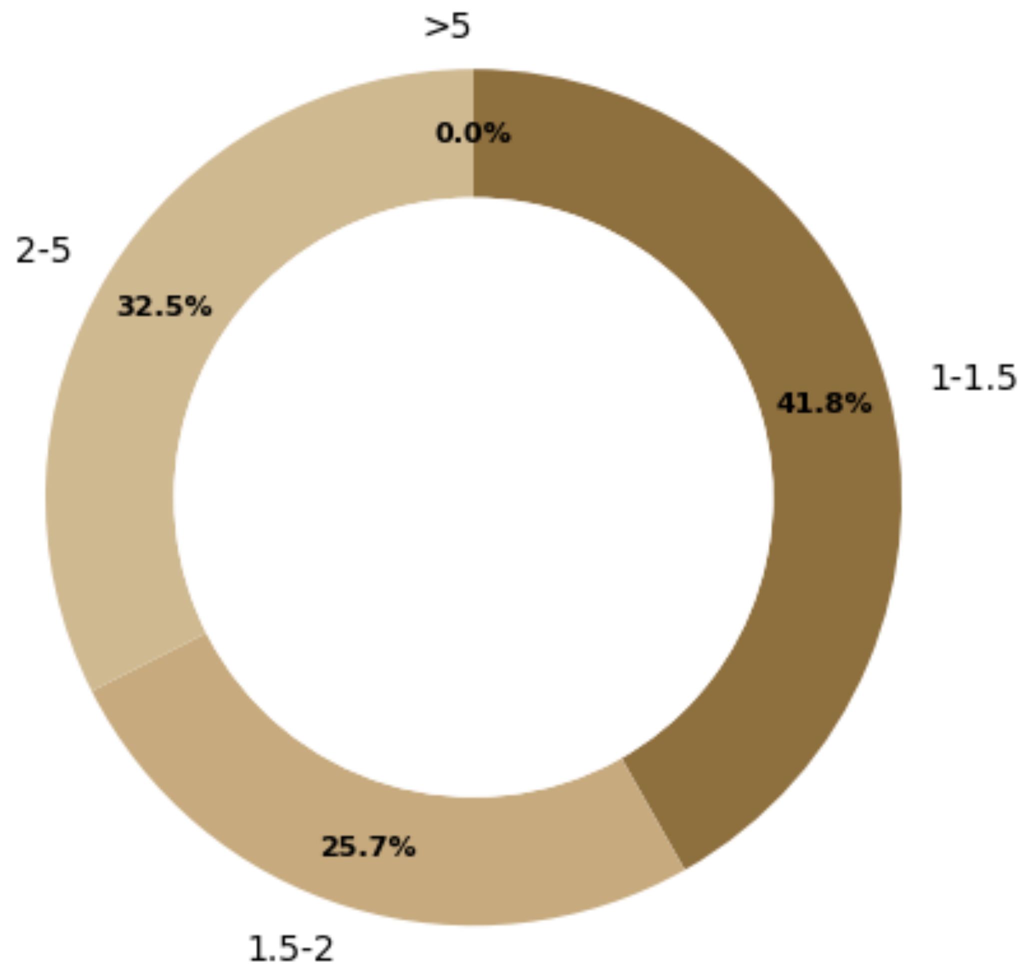
Samarth Bansal

Software Developer with 3 years of experience in Technology services

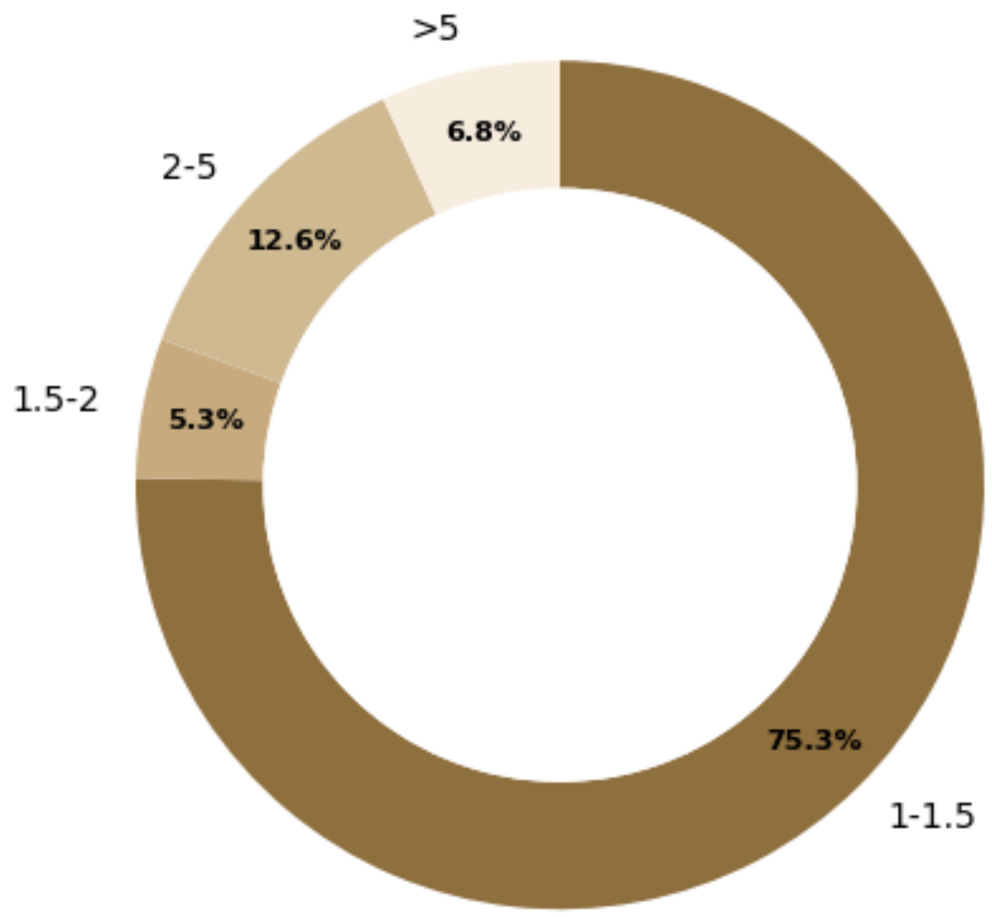
MAX FACTOR

Distribution of Max Factor shows a significant improvement as majority of materials have a max factor between 1-1.5

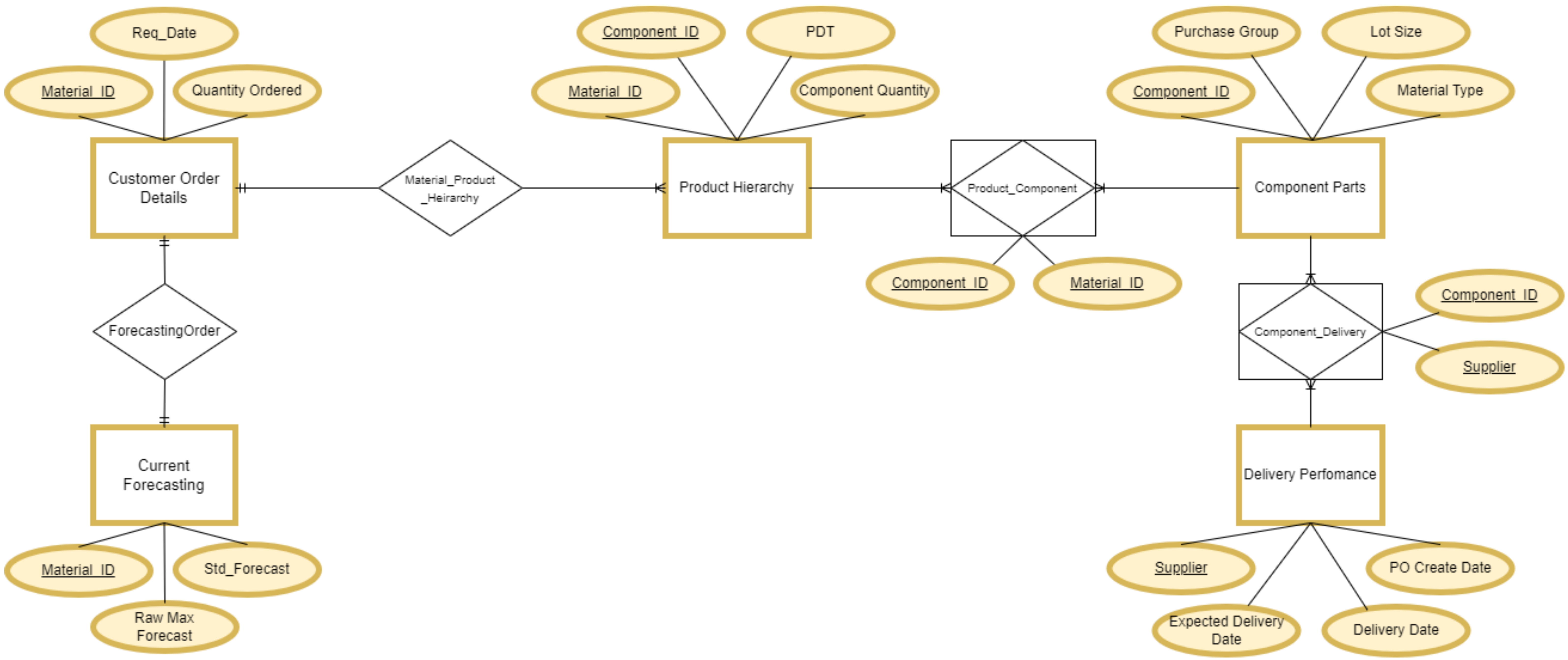
High volume



Low volume



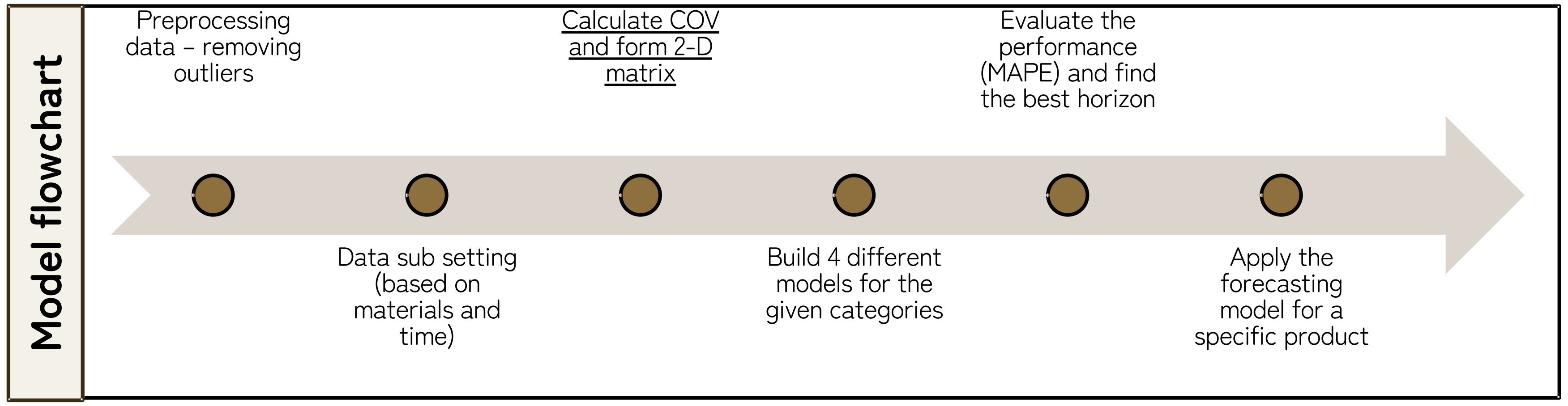
Entity Relationship Diagram



Models Applicability Check

Model group	Model Types	Use Case Description	Key Considerations
Traditional Statistical Time Series	• ARIMA/SARIMA • Holt-Winters • STL Decomposition	• ARIMA/SARIMA for time-dependent patterns in 'Qty Requested' • Holt-Winters for trends and seasonal effects • STL for analyzing seasonal components	• Best for clear trends or seasonality • STL for analysis, not forecasting • Limited use of complex categorical variables
Advanced and Flexible Forecasting	• Prophet ✓ • Machine Learning (RF, GB) • Deep Learning (RNN, LSTM)	• Prophet for daily patterns and irregular trends • Machine Learning for variable interactions • Deep Learning for large datasets with sequential dependencies	• Extensive data preprocessing • Computationally intensive • Ideal for nuanced demand patterns
Multivariate Time Series	• Vector Autoregression (VAR)	• Ideal for interrelated variables over time	• Assumes interdependence among variables • Requires consideration of variable relationships

Purch.Re(	Material	Delive	Qty Requeste	Req.Date	Deliv.dt	Setup	MM/	MRP C	MRP Ty	In House PRD Tir
35688872	K3562-2	0	2.000	12/13/2023	04/18/2024	0.00	0	870	ZW	10
35669309	K3562-2	0	4.000	12/11/2023	04/16/2024	0.00	0	870	ZW	10
35621330	K3562-2	0	1.000	12/05/2023	04/10/2024	0.00	0	870	ZW	10
35610567	K3562-2	0	9.000	12/04/2023	04/09/2024	0.00	0	870	ZW	10
35590946	K3562-2	0	10.000	12/01/2023	04/08/2024	0.00	0	870	ZW	10
35726700	K1500-2	0	1.000	12/18/2023	04/08/2024	0.00	0	875	ZW	8
35642386	K4630-2	0	1.000	12/07/2023	04/05/2024	0.00	0	875	ZW	15
35642450	K4352-1	0	9.000	12/07/2023	04/05/2024	0.00	0	875	ZW	15
35707964	K940-25	0	10.000	12/15/2023	04/05/2024	0.00	0	875	ZW	9
35707890	K163-1	0	6.000	12/15/2023	04/04/2024	0.00	0	875	ZW	8



Forecasting Iterations

Training period	Testing period
1st Jan- 2018 – 31st Dec 2022	1st Jan 2023- 31st Dec 2023
1st Jan- 2019 – 31st March 2023	1st April 2023 – 31st Dec 2024
1st Jan- 2020 – 31st March 2023	1st April 2023 – 31st Dec 2024

Data Transformation	Testing period
Scaling	Models trained with scaled features can more accurately capture the relationships between features and the target variable
Logarithmic transformation	Log transformation can help reduce right or left skewness, making the distribution more symmetric and closer to normal

Models tried	Reason
Prophet Grid Search	To perform hyperparameter tuning in order to enhance the model
AutoARIMA	To fit different values for p,d,q – according to the trend and seasonality in data
XGBoost Regressor	To avoid overfitting and sequentially learn from errors to provide better predictions
LSTM	LSTMs can learn and remember information over long sequences and are highly effective for tasks where the context or information from earlier in the sequence is vital for making predictions or decisions later on.
SARIMA – Seasonal ARIMA	To capture the seasonality present in the data

IT Infrastructure Overview

Software and Python Package Requirements

1 **The following software are required to refresh the statistical demand forecast model:**

python

- Execute Stat. Forecasting algorithms
- Create inputs for Tableau Dashboards

Tableau

- Create visual summary of forecast outputs
- Facilitate internal demand review discussions

Notepad ++

- Perform QCs for intermediate outputs
- Create formulae updates for python notebooks

2 **For the first model run, the following python packages from open-source repository are required:**

Package Name
pandas
matplotlib
scipy
prophet
…
fpp2
TTR
dplyr
MLmetrics

Package Name
numpy
scikit-learn
plotly
seaborn
…
reshape
reshape2
MASS
tidyr

Update in the approach

Last Week	This Week
Use lowest weighted MAPE to select best model	Use lowest weighted MAPE to select best model

Last Week

Use lowest weighted MAPE to select best model

- L. Var - L. Vol
  - Automated model selection
    - ARIMA
    - Holt Winter
    - Prophet
- L. Var - H. Vol
  - Automated model selection
    - CNN
    - Wavenet

This Week

Use lowest weighted MAPE to select best model

- L. Var - L. Vol
  - Automated model selection
    - ARIMA
    - Holt Winter
    - Prophet
- L. Var - H. Vol
  - Automated model selection
    - CNN
    - Wavenet
    - CNN Classifier + CNN
    - CNN Classifier + Wavenet

Why Classifier?

There were no order requests for most of the days

Req.Date	K1170	K1297	K1365-23	K1386-3	K1387-3	K1500-1	K1500-3	K1500-4	K1504-1	K1524-3	K1546-1	K1551-2	K162-1
1/1/21	96	0	0	0	0	0	0	0	0	0	0	0	0
1/2/21	0	0	0	0	0	0	20	0	0	0	0	0	0
1/4/21	12	24	0	0	0	30	0	0	0	0	0	0	0
1/5/21	96	48	0	0	0	0	0	0	0	0	0	0	0
1/6/21	108	24	0	1	0	0	0	0	0	0	0	0	0
1/7/21	0	36	0	0	1	0	0	0	0	0	0	0	0
1/8/21	96	0	0	0	0	0	0	10	0	0	0	0	0
1/9/21	0	0	0	0	0	0	0	0	0	0	0	0	0
1/11/21	0	24	0	0	0	30	0	0	0	0	0	0	30
1/12/21	96	36	0	0	0	0	0	0	0	0	0	0	0
1/13/21	96	12	0	0	0	0	0	10	0	0	0	0	0
1/14/21	96	12	0	0	0	0	0	0	0	0	0	0	30
1/15/21	96	0	0	0	0	0	0	0	0	0	0	0	0
1/16/21	0	0	0	0	0	0	0	0	0	0	0	0	0
1/18/21	96	0	0	0	0	0	0	0	0	6	0	0	0
1/19/21	96	72	6	0	0	30	0	10	0	0	0	0	0
1/20/21	96	12	0	0	0	0	0	0	0	0	0	0	0
1/21/21	96	12	0	0	0	0	0	0	24	0	0	0	0
1/22/21	12	0	0	0	0	0	0	0	0	0	0	0	30

Req.Date	K1622-1	K1622-3	K1690-1	K1702-1	K1726-5	K1728-6	K1737-1	K1745-1	K1759-70	K1759-95	K1770-1	K1780-3	K1803-1	K1803-2
1/1/21	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1/2/21	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1/4/21	0	0	0	0	2	0	0	0	0	0	0	0	0	0
1/5/21	0	0	0	0	1	0	0	0	0	0	0	1	0	20
1/6/21	0	0	0	0	9	0	0	10	0	0	5	0	0	0
1/7/21	21	0	0	0	2	0	0	0	17	0	0	3	0	0
1/8/21	0	0	0	0	1	0	0	0	0	0	0	0	0	20
1/9/21	0	0	0	0	1	0	0	0	0	0	0	0	0	0
1/11/21	0	0	0	0	3	0	0	0	0	0	0	3	0	0
1/12/21	0	0	0	6	4	0	0	0	0	0	0	0	0	0
1/13/21	0	0	0	0	2	0	0	0	0	0	0	0	0	0
1/14/21	0	0	0	0	2	0	0	0	15	0	0	0	0	0
1/15/21	0	20	0	0	2	0	0	0	0	0	5	1	0	0
1/16/21	0	0	0	0	0	0	0	40	0	0	32	0	0	0
1/18/21	21	0	0	0	5	0	0	20	0	0	0	1	0	0
1/19/21	0	0	0	6	2	0	0	30	0	0	37	0	0	0
1/20/21	0	0	0	0	4	0	0	0	0	0	0	0	0	0
1/21/21	0	0	0	0	7	0	0	0	0	0	0	0	0	20
1/22/21	24	0	0	0	1	0	0	0	0	0	0	5	0	0

Since the model is struggling to capture extreme crests and troughs, let's make it easy to at least identify troughs beforehand

"Classifier + Predictor" Model

- Training and Validation Data Preparation for model
- Use CNN classifier to identify the days which may have no orders and make them 0
- Apply CNN/Wavenet forecast model to the rest entries which may have order presence
- Compile both zero order days from classification and non-zero order days from prediction to build final forecast table